REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
14-04-2006	Final Report	July 2005 - July 2006

4. TITLE AND SUBTITLE

Analyzing the Utilization of Interferon-γ Screening at Recruit Training Command, Great Lakes

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Mitchell, David P. LT, MSC, USN

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Naval Hospital
3001A Sixth Street
Great Lakes, IL 60088-5230

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

US Army Medical Department Center and School
BLDG 2841 MCCS-HFB (Army-Baylor Program in Healthcare Admin
3151 Scott Road, Suite 1411
Fort Sam Houston, TX 78234-6135

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S) 32-06

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT The tuberculin skin test for diagnosis of *Mycobacterium tuberculosis* has many limitations when applied to serial testing of recruits in a restrictive training environment. One year ago the FDA approved QuantiFERON-TB Gold® as a blood test to screen for tuberculosis and latent tuberculosis. This product uses a combination of CFP-10 and ESAT-6 proteins to elicit in vitro response of white blood cells to detect prior exposure to *m. tuberculosis*. The main benefit of the test is it only requires one visit to the physician's office, saving both time and money. Because the test is new, a comparison of the two products singularly and in a two-tiered model was completed testing each product for accuracy, functionality, risk, and cost to determine which test was better suited for use in an environment where the protection of military readiness was the goal. Applying sensitivity and specificity values from journal publications; it was found that two-tier testing was most accurate and cost-effective, while the tuberculin skin test was the most risk averse, and the QuantiFERON-TB Gold had the highest process functionality. When applied to a decision matrix based upon the priorities of the military, the conclusion is to continue to use the Mantoux TST due to the lower overall risk to military readiness.

15. SUBJECT TERMS

Tuberculosis, tuberculins skin test, QuantiFERON-TB Gold®, ESAT-6, CFP-10, interferon-gamma testing, military readiness, Navy Recruit Training.

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE	SAR	41	Ms. Renee Pryor
					19b. TELEPHONE NUMBER *(include area code)* (210) 221-6443

Running Head: Interferon-γ screening for TB

Graduate Management Project:

Analyzing the Utilization of Interferon-γ Screening for Tuberculosis at Recruit Training Command, Great Lakes

Presented to MAJ. M. Patrick, MHA, PhD.

By LT David P. Mitchell

Army-Baylor Health Administration Program

Ft. Sam Houston, TX

May 31, 2006

Acknowledgements

I would like to sincerely thank the many people who helped me through this project. The first of these is Captain Rebecca McCormick-Boyle, my residency preceptor at Naval Hospital, Great Lakes. She introduced me to the project through a business case analysis conducted on this same subject. CAPT Monestersky, Preventive Medicine; LCDR Jacobs, Occupational Medicine; Mr. Lesko, Industrial Hygiene; and the preventive medicine staff at USS Tranquility also contributed significantly in the completion of this analysis. Finally, I would like to thank my wife Susan and son Tyler for allowing me the time to complete the project and supporting me whole-heartedly in my pursuits.

Abstract

The tuberculin skin test for diagnosis of *Mycobacterium tuberculosis* has many limitations when applied to serial testing of recruits in a restrictive training environment. One year ago the FDA approved QuantiFERON®-TB Gold as a blood test to screen for tuberculosis and latent tuberculosis. This product uses a combination of CFP-10 and ESAT-6 proteins to elicit in vitro response of white blood cells to detect prior exposure to *m. tuberculosis*. The main benefit of the test is that it only requires one visit to the physician's office, saving both time and money. Because the test is new, a comparison of the two products singularly and in a two-tiered model was completed testing each product for accuracy, functionality, risk (both patient and government), and cost to determine which test was better suited for use in an environment where the protection of military readiness was the goal. Applying sensitivity and specificity values from journal publications, it was found that two-tier testing was most accurate and cost-effective, while the tuberculin skin test was the most risk averse, and the QuantiFERON®-TB Gold testing process was easiest for staff and patients. When applied to a decision matrix based upon military priorities, it is recommended that RTC Great Lakes utilize the Mantoux TST followed by the QuantiFERON®-TB Gold in a two-tier test process.

Table of Contents

List of Tables

List of Figures

Introduction

Tuberculosis (TB) has been called phthisis, wasting disease, Pott's disease, lupus vulgaris, consumption and the White plague. It is the leading cause of death from an infectious agent throughout the world with over 1.8 million deaths attributed to it annually (World Health Organization, 2004). The infecting agent, *Mycobacterium tuberculosis*, has been killing humans for thousands of years and remains one of the largest public health problems in the world with 19% to 43% of the earth's population being infected (Dunlap, Bass, Fujiwara, Hopewell, Horsburgh, Salfinger, et al., 2000). The United States and other developed countries have significantly lower prevalence rates than second and third world countries. Nevertheless, it remains a legitimate threat to the health of the U.S. populace with an estimated three to five percent being infected with *Mycobacterium tuberculosis* (Bass, 2003). Treatments are available for those infected with active TB, although some strains are drug resistant. The most recommended treatment is to identify those who have a latent tuberculosis infection (LTBI) and treat them prophylactically with isoniazid before they exhibit symptoms of the disease.

Mycobacterium tuberculosis is extremely slow growing, which presents problems for scientists and public health officials alike. Some newly infected persons will very quickly succumb to the disease and display symptoms. These people need to be identified and treated rapidly to forego spreading the disease. Others undergo a long latency period where they are carriers but will never exhibit symptoms of the disease. Somewhere between these two extremes are the five to ten percent of individuals who will succumb to the disease, develop symptoms and progress to active TB. These individuals are the vectors who spread the disease to others (Dunlap et al. 2000).

Acid fast bacillus smear microscopy and the Mantoux Tuberculin Skin Test (TST) are the most commonly used tests for the diagnosis of TB and LTBI (Taylor, Nolan, & Blumberg,

2005). They are based on science founded in the 19th century. The current Mantoux TST is a test that identifies the latent carriers before they can spread the disease. Unfortunately, no means exists to predict how long it takes for a person to react to the TST after they have been infected. Also, once a person has an active infection, there is an increased possibility they will not react to the TST compared to those with LTBI (World Health Organization, 2005).

Conditions that Prompted the Study

Because of the high prevalence and significant public health threat, there has been an extensive effort to reduce the spread of TB throughout the world. The United States has one of the most successful TB elimination programs in the world. This has been accomplished almost exclusively by using the Mantoux TST for LTBI and active TB detection. Consequently, after 70 years of use, the mortality and prevalence rates are extremely low in the U.S. However, there is still not a definitive "gold standard" for the diagnosis of LTBI or active TB. The Mantoux TST is widely criticized for its lack of specificity, inter-operator error and indeterminate sensitivity (Dunlap et al, 2000). Consequently, its positive predictive value (the probability that a patient with a positive test truly has the disease) is estimated to be as low as 16% in some segments of the U.S. population (Bass, 2003). This accounts to 84% of the people who are placed on medication to rid them self of LTBI as false positive. They are treated unnecessarily, at great costs to their insurance payer or themselves. This practice is also putting them at some risk for reactions and side effects from unnecessary medication.

The ability of an infectious disease, such as TB, to reduce military readiness has prompted the Navy to develop an extensive TB surveillance program. The program starts when all enlistees enter the Navy, at Recruit Training Command (RTC), Great Lakes. At RTC alone, the Navy has averaged 40,000 Mantoux TSTs and treats an average of 1,900 recruits for LTBI annually. The proclivity of having numerous people living in close quarters, such as on a ship or

in a field environment, can precipitate cross-contamination and lead to a rapid and widespread outbreak (Smith, & Schillaci, 1987). Closed ventilation systems used on ships increases the likelihood of an outbreak (Lamar & Malakooti, 2003). In order to minimize this, TB screening starts within the first few days of entering Recruit Training (RT) and is continued throughout the careers of today's sailors.

RT is better known as boot camp, where young men and women make the transformation from civilian life to that of a sailor. During their eight weeks at RT, recruits undergo 16 hour days of formal training in subjects such as the Uniform Code of Military Justice, basic seamanship, naval history and naval organization. The enlistees' schedule is full from before sunrise to long after sunset with training, tests and inspections.

An important aspect of the accession processing at RTC is the medical care. Within the inprocessing clinic, Branch Health Clinic (BHC) 1523, recruits are screened for dental diseases, visual acuity, human immune deficiency virus (HIV), glucose-6-phosphate dehydrogenase deficiency (G6PD) and TB before beginning formal recruit training. They are also immunized against communicable diseases including meningococcal; measles, mumps and rubella; hepatitis A and B; tetanus and diphtheria; polio; varicella; yellow fever; and influenza (seasonally from October through May). These initial screenings and tests are done on "P" days. "P" days are so named as they do not count towards the eight weeks of boot camp. They are preliminary days designated for completion of administration matters, hair cuts, uniform issue and preventative medical care. It is important for these functions to be accomplished within the first few days upon arrival at RTC so the Navy can identify individuals who may not be suitable for service. For the majority who are suitable, the "P" days allow time to complete necessary business that goes along with entering the military while minimizing disruption to the training cycle.

Thirty days after the first round of immunizations, recruits are marched back to the inprocessing clinic to undergo their second round of immunizations which includes hepatitis A and B and the second varicella vaccination. Otherwise, all further medical care while in boot camp is conducted at Branch Health Clinic 1007 (BHC 1007).

In years past, the training cycle was more flexible at RTC which allowed health care workers more access to recruits. Efforts to reduce costs have led to the review of every facet of training as well as every hour scheduled during the eight week training program. One result of this was the elimination of service week. Service week was a time where no training was conducted. Instead, recruits worked in the galley to prepare food for the other recruits. This week was traditionally used by the medical staff to follow-up with recruits for extensive dental or medical treatment in an effort to reduce the time away from training. An untoward consequence of its elimination has been the reduced access for recruit medical care.

The clinical processes at BHC 1007 have adjusted to the tighter schedule. This clinic is the primary care clinic for recruits and contains specialty care for orthopedics and mental health. It is also where special physical exams and screenings for submarine or sea duty, divers, special operations, and overseas screenings take place. The recruit population will exceed 10,000 during some months of the year, which significantly stresses the resources in place to take care of them. The high volume of recruits coupled with the very tight schedule makes for a challenging environment for medical care, especially preventative health programs, such as the TB surveillance program.

Civilian standards for "booster" TB testing cannot be utilized here. This is in part due to the tight time schedule, but it is mostly due to the stress on the recruits immune system caused by seven or eight immunizations that recruits receive within the first three days upon arrival. The cumulative affect of these immunizations has been shown to weaken the body's response to the

TST. To overcome the omission of the booster test, the Navy has decreased the positive cut-point for Mantoux TST testing down to 10 mm. In a low risk civilian population, a positive test result occurs at 15 mm. Clinical personnel also conduct a chest radiograph and interview any recruit with a test that ranges from 5 to 9 mm to ensure there are no signs of active or latent TB infection. This is done to ensure that no one with active TB or LTBI slips through undetected.

Efforts to reduce time in training and overall costs at RTC have led to a review of all medical screenings. QuantiFERON®-TB Gold is a new blood test used for tuberculosis screening. It is based upon interferon-γ response and has been approved for use in the U.S. by the Food and Drug Administration (FDA) in December 2004. QuantiFERON®-TB Gold (QFTG) is manufactured and marketed by Celestis Incorporated (Victoria, Australia) as the "gold standard" for *Mycobacterium tuberculosis* screening. It is based on ELISA[1] methodology and claims to be more specific than the Mantoux TST (Celestis Product Information Insert, 2005). The main advantage of a blood test over the in-vivo test is the elimination of the follow-up visit required for the interpretation of the TST. Blood samples are already taken from the recruits for the HIV, hepatitis, and G6PD tests. The addition of another blood tube added to the blood draw in this process would not add significant time or costs to the existing venipuncture process, easily outweighing the time and cost resources exhausted with the two visits necessary to conduct the Mantoux TST.

The research problem addressed in this analysis questions whether QFTG test is better, equal or worse than the Mantoux TST in identifying individuals who have active and latent *Mycobacterium tuberculosis* infections. A better screening test would provide faster identification and treatment of those afflicted. It would save time for the recruits and medical

---

[1] ELISA (Enzyme-linked Immunosorbent Assay) is a test method to determine the immunological response of antibodies in blood or urine to a specific protein, peptide or drug.

providers alike by reducing unnecessary treatment for those who have historically been identified as false positive, thus reducing the impact on public health resources. The costs of each test scenario will also be calculated in an effort to compare the cost of each program, including those costs associated with treating individuals who are false positives, as well as the potential costs of overlooking the false negative populations at RTC. Finally, government risk will be analyzed in each scenario to determine which test is best at reducing the risk to military readiness. Military readiness is based upon and individual's physical ability to complete their assigned job. This cannot be jeopardized because of a faulty test; that is an unacceptable risk. Within the test scenarios, the biggest risk to government readiness is the misdiagnosis of an active case of TB which could lead to a TB outbreak. This would cost hundreds of thousands of dollars and potentially impact the capability of a war-fighting unit. This is a critical concern in a military analysis because increasing the risk by changing the test or process is likely to be met with great resistance for fear it would decrease force readiness.

Literature Review

Frequent traveling and long deployments to foreign countries place military personnel at a higher risk for TB than the general population (Lamar et al. 2003), as well as potentiate the spread of communicable disease. An example of this was the Influenza outbreak of 1918. That pandemic is reported to have begun in Fort Riley, KS, and then spread globally by traveling servicemen in World War I. In the case of tuberculosis, service members are likely to be exposed overseas in countries with a high prevalence of TB, and then contaminate other military members in their unit or civilians in the U.S. upon their return. An example of this occurred in 1998, when a Marine assigned to a Navy ship was not prescribed pharmacotherapy after a TST conversion was erroneously classified as negative. In this instance the induration was attributed to improper TST administration technique. The Marine subsequently became ill, and after

repeated trips to the ships' sick call was discovered to have active tuberculosis. An epidemiological outbreak investigation later determined that the Marine was the source of 712 latent TB infections, as well as 21 active TB cases (LaMar et al.). The probability of an outbreak attributable to misinterpreted results needs to be reduced, especially at RTC, where thousands of Sailors could become infected if an active case was not properly identified upon a persons' entry into the Navy.

There are dozens of articles that research the validity of complaints against the Mantoux TST. The American Thoracic Society in conjunction with the CDC publishes guidance for medical professionals on this periodically. Their most recent guidance, *Diagnostic Standards and Classification of Tuberculosis in Adults in Children* (Dunlap et al. 2000) specifies a list of the factors that cause false negative tests. These factors include infections from diseases such as measles, mumps, chicken pox, HIV, typhoid, brucellosis, typhus, leprosy, pertussis, overwhelming tuberculosis and South American blastomycosis. Vaccinations, poor nutrition, corticosteroid use, age and stress are also listed as causes of some false negative tests. Controllable factors such as the handling of the PPD, its exposure to light and extreme temperatures, how long its been in the syringe, how it was administered, and the method in which it was read have also been listed as factors which can affect the outcome of the test (Dunlap).

Two main causes have been identified for the majority of false-positive tests. One is from persons who have previously been administered a BCG[2] vaccination (Jensen, Lambert, Iademarco, & Ridzon, 2005). This is not a concern for the majority of the U.S. population because BCG is not recommended for use here. The Navy enlisted population however does

[2] BCG (Bacillus Calmette-Guerin) vaccine is used extensively in high prevalence countries to prevent *Mycobacterium tuberculosis* infection in children but is not recommended for use in the U.S.

have an increasing number of foreign born recruits. In 1992, the percentage of foreign born recruits had risen to over 9% (Smith, Ryan, Gray, Polonsky & Trump, 2002), and data from Mazurek (2004) indicates the number is closer to 10%. This may be cause for concern when evaluating the TST program and any future changes to it.

The other significant cause of false-positive tests is exposure to environmental mycobacterium, also known as non-tuberculin mycobacterium (NTM). These bacterium are of the same genus as *Mycobacterium tuberculosis*, but do not infect humans. Geographically, the southern and eastern portions of the U.S. have a higher rate of environmental mycobacterium (Dunlap et al. 2000). But within regions, it is hard to determine the overall effect of NTM and the specific impact on TST testing is difficult to quantify (Bass, 2003).

The two visit process that requires the patient to make a return visit to the health care provider from 48 to 72 hours after receiving the PPD injection is another problem with the Mantoux TST. The need to have the subject's arm "read" results in a substantial number of tested patients failing to return, leaving health care providers with unknown results. The cost of conducting a test without the benefit of getting any results is a waste of health care resources.

The accuracy of the reading has also been a source of concern in some studies. Kendig et al. (1998) reported that over 90% of TSTs read by health professionals were incorrectly interpreted. This has resulted in the research and discovery more effective ways to read the test, including ultrasonic measurement. Recommendations from these studies include having only physicians read TSTs, as well as emphasize attention to detail for the entire skin testing process. The basis of the PPD test was termed "relatively crude" (Bass, p. B-3, 2003) because of the in vivo results being dependent upon a t-cell response from a protein mixture that is over 70 years old.

Another area of concern with the Mantoux TST is the extent to which a person will react to the tuberculin protein after long time lapses between the exposure and the PPD test. Persons who have received a Mantoux TST for the first time in many years and return with a reaction between 5 and 14 mms are recommended for two-step "boosting". Boosting is a follow-up test given one to three weeks after the implantation of the first test. It is used to determine if the patient has a new LTBI infection or if they have an old infection that their immune system did not react to because of the length of time between the initial exposure and the test (Dunlap, Bass, Fujiwara, Hopewell, Horsburgh, Salfinger, et al. 2000). Current guidelines recommend that persons who undergo serial TB testing start off with a two-step test as their baseline result. All subsequent tests will then be based upon an increase of 10 mm or more induration to determine if it is a new infection. The goal of two-step testing is to decrease the probability that a person will be identified as a new converter, when in fact they have an old infection. The benefit of this method is it will reduce the resources devoted to epidemiological outbreak investigations, which could be erroneously initiated if one year after their initial PPD they boosted a reaction from an old infection.

A 2005 study in northern California demonstrated a correlation between Mantoux TST boosting and a positive QFTG (Nguyen, Perry & Parsonnet, 2005). For this study, participants were initially tested with both the Mantoux TST and QFTG and then re-tested with the TST three months later. Results of the first tests noted a discordance of 16 QFTG positive patients who were TST negative. After the second TST, this discordance was reduced by half leading the authors to conclude that QFTG will frequently anticipate a TST boost.

The need for a more effective test for LTBI as well as for active TB has been cited for many years. The Institute of Medicine (2000) report, *Ending Neglect: The Elimination of Tuberculosis in the United States,* listed this as one of its priorities for 2001. The World Health

Organization has also prioritized the detection and treatment of tuberculosis. The QuantiFERON®-TB Gold test has the potential to assist each of these organizations. The Center for Disease Control (CDC) published in their Mortality and Morbidity Weekly report that QFTG is an equal test to the TST for identifying TB and LTBI (Mazurek, Jereb, LoBue, Iademarco, Metchcock, et al. 2005). This test is based upon interferon-γ response to proteins in whole blood and has been touted as a better way to screen for TB since 1998. Clinical research trials of interferon-γ test have been published in numerous journals. Reports prior to 2002, however were based on a product that used PPD based interferon-y response. These tests were prone to the same false results seen in the Mantoux TST that were earlier discussed. In 2002, the test switched from using PPD to illicit the cytokine response, to the proteins ESAT-6[3] and CFP-10[4]. These are the proteins used in the QuantiFERON®-TB Gold product. Upon the FDA approval of QFTG, the first generation product, QuantiFERON-TB, was removed from the market. Because of its removal, only studies using antigens from the proteins ESAT-6 and CFP-10 will be cited in this paper.

Results of the interferon-γ studies vary. This in itself is not unusual and can be attributed to studies being done in different countries with different populations and different TB prevalence rates and the utilization of different study designs (Whalen, 2005).

Ferrara et al. (2005) concluded that in BCG vaccinated populations, QFTG has an increased specificity compared to the TST. This study was conducted in an Italian hospital with nine patients enrolled who were diagnosed with TB. Of these nine, QFTG correctly identified six of them compared to only three who were identified with the TST. It also reported one patient who tested QFTG negative whom later was diagnosed with TB. This incident and the issue of

---

[3] ESAT-6 (early secretory antigenic target 6)
[4] CFP-10m (culture filtrate protein 10)

having a high number (68 of 318) of indeterminate results are attributed to the large number of patients (65) that were undergoing therapies known to be immunosuppressive.

Kang et al. (2004) conducted a side by side comparison trial of TST to QFTG in Korea. For their research, they divided their sample into four risk groups. Group one was made up of persons having no known risk factors for TB, group two had a slight risk, group three had a moderate risk, and group four was confirmed TB. In all groups, QFTG identified from 47% (group 1) to 27% (group 3) less positive tests than the TST. And of the 54 cases that were diagnosed as active TB, 46 (85%) were QFTG positive. The researchers concluded that QFTG was more specific in areas of high tuberculosis prevalence where the population is mostly BCG vaccinated.

Mori et al. (2004) conducted a study in Japan based on two population groups, both of which were BCG vaccinated. One group had culture confirmed TB and the control group of nursing students had no risk factors for TB. This study resulted in a specificity of 98.1% for the QFTG as compared to 68.1% for the TST. Again, QFTG demonstrated higher specificity when compared to the TST in a BCG vaccinated population.

Ravn et al. (2004) conducted a prospective study from a small cohort in Denmark of patients who were suspected of having TB. This research was conducted to see how QFTG would identify patients with active TB when compared to microscopy and sputum smears. In this test, the 73 who were screened with QFTG resulted in 51 who were positive and 22 who were negative. Of the 51 positives, 41 (80%) were determined to have active TB from microscopy exams and sputum smears. From the ten which were designated false positives, eight had risk factors for LTBI including three with HIV.

Brock, Weldingh, Lillebaek, Follmann and Andersen (2004) utilized both the QFTG and the Mantoux TST in an outbreak investigation in a Danish high school, comparing the two tests

side by side. Of the 700 contacts, 37 were TST positive using a two tuberculin unit Mantoux TST (>10 mm was positive cut-off). One hundred twenty-five from the group of 700 volunteered to be in the QFTG experimental group. Eighty-five out of the 125 were not BCG vaccinated, 40 were BCG vaccinated. In all, 29 of the 37 students with a positive TST also underwent the QFTG. Of the 29 positive TSTs, 27 (93%) were also QFTG positive. Furthermore, one out of the remaining 97 was positive for QFTG that was not TST positive. Since all patients who were either QFTG positive or TST positive were offered isoniazid treatment and there is no definitive standard for diagnosis of LTBI, it is not known which test performed better.

These studies demonstrate promising results utilizing the QFTG test. The problem with correlating their results to a Navy enlisted population is the sample is not representative of what would be expected in the U.S. The biggest concern is their BCG history. Another concern is the use of a different TST protein in a smaller (2 TU) dose. These variables lead to dissimilarities that are difficult to account for. The U.S. Navy recruit cohort is not a population that is likely to be BCG vaccinated and thus diminishes the utility of the QFTG studies. This leads to great difficulty when comparing the Mantoux TST to the QFTG in a U.S. sample.

Fortunately there was one study that was conducted at Naval Training Center Station, Great Lakes. This study, conducted by Mazurek, Zajdowicz, Hankinson, Costigan, Rothel, Toney, et al., compared the QFTG product to the Mantoux TST on over 800 new enlistees in 2004. It is not yet published, but compares the Mantoux TST to the QFTG with a population demographic that is the same as the sample. Their study uses a prospective analysis comparing the results of a Mantoux TST to that of the QuantiFERON®-TB Gold first generation product as well as that of QuantiFERON®-TB Gold-TB Gold®. When comparing the results of 828 subjects who participated in the TST injection and reading as well as the QFTG test, the specificity of the QFTG was 99.8% as compared to 99.1% for the TST when using a 15 mm cutoff for the TST.

The agreement between the two tests was 95.7%, also based on the 15 mm TST cutoff. This cohort is much more representative than the other studies which have tested QFTG to date. Because of this, it will be frequently cited in throughout the rest of this analysis.

Other promising results from this study are the lower incidence of indeterminant results from the QFTG test in the sample population. While other studies have indeterminate results ranging from 5% to 20%, this study was only at 2%. The lower result as a part of this study dramatically increases its utility for a test that is supposed to reduce the patient visits from one to two in the new process.

There are some concerns with test discordance in the Mazurek study. There were no instances of discordance resulting from a positive QFTG and a negative TST. There was significant discordance, however, from positive TSTs and negative QFTG tests. Of the 19 positive TSTs between 10 and 14 mms, there were no positive QFTG's. Also, of the 19 positive TSTs that were greater than 15 mm, there were only five positive QFTG's. 22

Petitti (2000) described numerous methods for determining cost-effectiveness when trying to establish proper values from numerous scientific studies. These methods can than be used for decision making when comparing different tests or treatment options. The meta-analysis technique she described will be used to assist in compiling the separate study results into a single number that can be used for test efficacy calculations.

Purpose

The purpose of this report is to determine if there is a difference between the Mantoux TST and the QuantiFERON®-TB Gold tests in identifying individuals with LTBI as well as active tuberculosis. This analysis is driven by results from 2x2 epidemiology tables developed using the characteristics of the Mantoux TST and the QuantiFERON®-TB Gold tests. These characteristics were then applied to estimated RTC Great Lakes future recruit populations using

historical LTBI prevalence, as well as historical treatment rates to project the outcomes if each test were used. The null hypothesis is there will be no difference in outcomes between the two tests. If it is determined that there is a difference in the ability of the QFTG to identify carriers of TB and LTBI, the null hypothesis will be rejected and further models will be constructed to determine cost, risk, accuracy and the utility of the two tests when used in the constrained time schedule that exists at RTC. Included in the modeling will be a two-tier testing scenario, where the TST is completed first and if it returns a positive result, a QFTG test would next be conducted.

Methods and Procedures

To evaluate which test has the greatest clinical accuracy and is most likely to correctly identify carriers of LTBI as well as identify those who are not, values were calculated for TB prevalence at RTC and the operating characteristics of each test. The sensitivity and specificity of the Mantoux test is not absolute or well documented, so estimations were used to project test results.

For the TB prevalence rate, the CDC defines prevalence as the number of positive converters in a TB screening program (CDC, 1997). To determine this, data from the Annual Summary of Tuberculosis reports was gathered for the past three years at RTC Great Lakes. These reports indicate an average prevalence rate of 5% for the recruit population. Table 1 lists this data from the Annual Summary reports. The average number of reactors per year is calculated to be 1,960, which was also used in other calculations.

The concern from using a flat 5% prevalence rate is that it does not compensate for false positive results when constructing epidemiology tables. Bass (2001) contends that the positive predictive value for the TST is 83% when the prevalence is 5%. Given the historical 5% prevalence at RTC and applying the 83rd percentile to it, the 5% prevalence leads to an adjusted

Table 1

Prevalence of LTBI at RTC Great Lakes

Year	Number Tested	Number of Reactors	Prevalence
2004	38938	1599	4.11%
2003	37858	2558	6.76%
2002	39843	1723	4.32%
Total	116639	5880	5.04%

prevalence of 4.15%. This number is between the 3 to 5% estimate provided by Taylor et al. (2000), and was used for all prevalence calculations that will be conducted in the course of this analysis.

Bass (2001) also has written that the likely sensitivity of the Mantoux TST in the United States is 95%. This is a highly debatable number, but without a gold standard for identification of LTBI, it cannot be empirically confirmed. It must be stated that the sensitivity of a test decreases as the prevalence in a population decreases. This means within the U.S. population the test characteristics may vary. With the prevalence at RTC being higher than rates in the general population, the 95% sensitivity was used for TST calculations throughout this study, even though it is only estimated.

To determine the specificity of the Mantoux TST, the estimates published by Bass (2001) was again used. He has written that the sensitivity of the TST is between 99% and 99.5%, again within the general U.S. population. RTC has a 10% foreign born population, of which from 2% to 9% have been BCG vaccinated (Mazurek, unpublished). This vaccinated sub-population will be prone to false positive results (Mori et al. 2005, Kang et al. 2004). As a result of this, the lowest specificity estimate of 99% was used for the TST calculations.

The sensitivity and specificity for QuantiFERON®-TB Gold will be easier to determine as this product has been compared to the TST in numerous recent studies. In those studies, QFTG results have been compared to Mantoux TST results, with either the Mantoux TST or the diagnosis of active TB as the discriminating factor of comparison. Because of the comparison to TST and the difficulty in identifying *Mycobacterium tuberculosis* in culture or microscopic exam, the true characteristics of the test may not be known. Currently there are no better tests, so the direct comparison is difficult to quantify. A meta-analysis of test sensitivity from study results are listed in table 2.

Table 2

Meta-analysis of QFTG sensitivity from published studies.

Study	Active TB	Pos QFTG	Neg QFTG	Sensitivity
Ravn et al	48	40	8	83.33%
Mori et al	118	105	13	88.98%
Kang et al	54	44	10	81.48%
Total	220	189	31	85.91%

The information from the above studies is insightful, but does not lend itself to a direct application in a U.S. population sample. The Mazurek study did not report sensitivity, but the data indicates it to be 85% which is in line with the foreign studies. The specificity reported by Mazurek was 99.8%, although this is harder to compare because of the differences in the manner that each study was set up. Because of the similar sample when compared to the projected sample, the Mazurek findings were used for all QFTG calculations within this report.

Using the stated prevalence within RTC, as well as the sensitivity and specificity for each test, Bayes' Theorem can now be calculated to determine the positive predictive value and negative predictive value of each test within the recruit population (Jekel, Katz, & Elmore,

2001). The results of the predictive values will be rated in the decision analysis as the measure of government risk. The higher the predictive value leads to more accurate identification persons who may develop active TB. If they go untreated, the risk of a TB outbreak will increase. This could lead to reduced force readiness as well as a substantial expense.

Patient risk was also considered in the final decision matrix. The patient does have a risk if they are not properly identified as having LTBI or TB; however, this risk would be the same if no test were done at all. There is also patient risk if they are placed on nine months of INH therapy. For the age of the sample population however, the risk is estimated to be less than 1 in 100,000 (Taylor et al., 2000). Therefore, the patient risk was not added as a component of the decision matrix and only the government risk will be used.

To determine the number of recruits involved in the tuberculosis surveillance program, 2x2 epidemiology tables were constructed. The characteristics of each test (sensitivity and specificity) are then applied to the RTC estimates of 37,500 projected recruits for years 2007 through 2009. Applying those results to the 2x2 tables, the number of positive, false positive, negative and false negative persons were calculated. The number of true positives and true negatives from each 2x2 table are then compared to determine which test scenario has the highest overall accuracy. This will be calculated by simply adding the two results.

Next, these figures provided the basis for applying costs to each program in an evaluation conducted to determine which test is most cost effective. The number of recruits who are projected to test positive within each scenario will be applied to the cost of treating those positive recruits. Costs to be considered for each program included the cost of the test, x-rays, pharmaceuticals and clinic visits for the population who is identified positive for each test. Another factor which was built into the cost calculations is the number of recruits who have historically been treated based upon their history alone. For RTC, this number is 156. The

patient's history is one of the best indicators in determining if a patient is prescribed INH prophylactically (Taylor et al., 2000). To ensure the accuracy of all the tables and processes, 156 patients will be taken off of the recruit sample size and built into the cost of treatment for each test process.

Costs associated for each of the screening and treatment steps were projected Naval Hospital MEPRS and EAS IV computer systems data. These costs are the best determination of total costs per visit, per x-ray, per prescription, as well as the cost for each screening test. The total cost of the separate programs will then be compared to determine the most effective method for screening and treating LTBI at the RTC clinics. It will be added to the decision matrix to evaluate the financial impact of each testing scenario.

Results

The positive predictive value of the Mantoux TST given a prevalence of 4.15, a sensitivity of 95% and a specificity of 99%, is calculated to be 80.44% using the formula below.

$$\text{Positive Predictive Value} = \frac{.95 \times .0415}{.95 \times .0415 + (1-.99) \times (1-.0415)} = .8044 \quad (1)$$

These results mean that the probability of a person with an abnormal test truly has the disease is 80%, and likewise over 20% of all positive results are incorrect. Using the same values, the negative predictive value is 99.78%. This indicates that the test is very accurate when identifying those persons without TB or LTBI with the sample population and stated prevalence. The calculation for the negative predictive value is shown in the formula below.

$$\text{Negative Predictive Value} = \frac{.99 \times (1 - .0415)}{(1 - .95) \times .0415 + .99 \times (1 - .0415)} = .9978 \quad (2)$$

Using the same prevalence with a sensitivity of 85% and a specificity of 99.8%, the positive predictive value of the QFTG is calculated at 94.85%; while the negative predictive value is 99.3%. This demonstrates a dramatic difference in the positive predictive value favoring the QFTG, although the negative predictive values are virtually the same.

$$\text{QFTG Positive Predictive Value} = \frac{.85 \times .0415}{.85 \times .0415 + (1-.998) \times (1-.0415)} = .9485 \quad (3)$$

$$\text{QFTG Negative Predictive Value} = \frac{.85 \times .0415}{(1 - .85) \times .0415 + .998 \times (1 - .0415)} = .993 \quad (4)$$

Figure 1 has been included to show the difference in the positive predictive value of each test from a prevalence of 20%, representing the foreign born recruits, then adjusted downward ranging from 4.5% to 1%. This was done to demonstrate the impact of disease prevalence on a test's effectiveness.

Also included in Figure 1 is a two-tiered test process designed to decrease costs of using the QFTG. This cost savings is realized by testing all recruits first with a TST. Those recruits whose TST reads greater than 5 mm would then undergo a QFTG test as a confirmatory test before treatment begins. Other serial TB testing programs would complete booster TSTs at this point. However, the recruits are scheduled for their first round of immunizations during this same visit. Currently, providers order x-rays, examine, and interview the 5 mm to 9mm reactors just as they would the 10 mm plus reactors to determine if they have active TB or LTBI. Many go untreated, and because of the scheduling demands at RTC, are not further evaluated until they report to their next command. By utilizing the QFTG at this juncture, the physician would have another piece of information to help determine whether or not to treat the recruit. The increased

accuracy of this test without having to bear the full cost of using it on every recruit is an option that will be explored in the other testing comparisons. As the figure indicates, the two-tiered test greatly exceeds the positive predictive values of the other two tests.

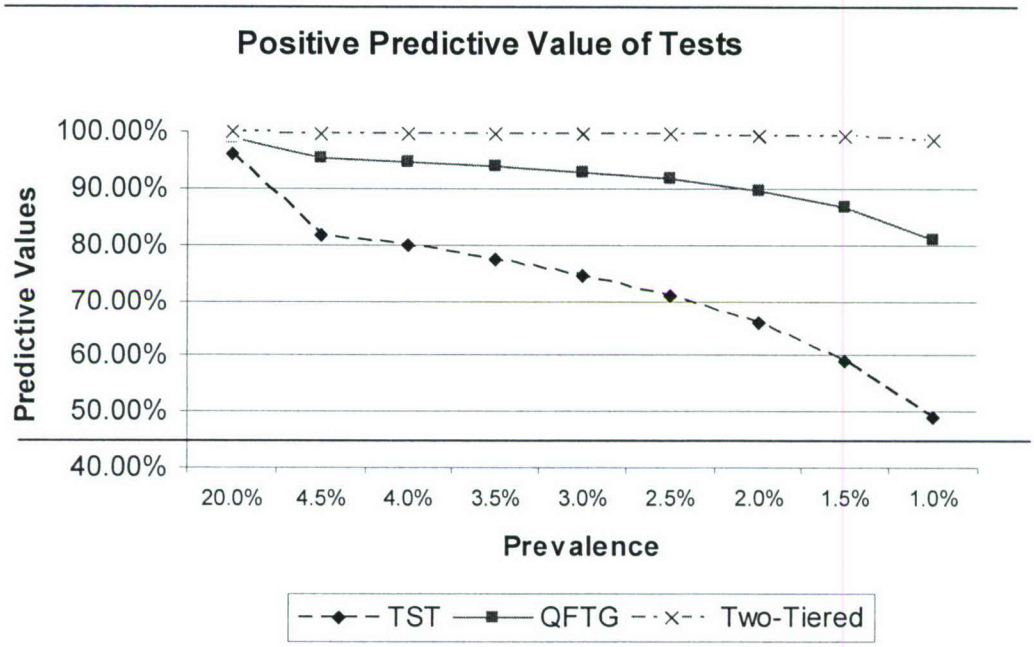

Figure 1. *Positive Predictive values of the TST, QFTG, and a two-tiered test.*

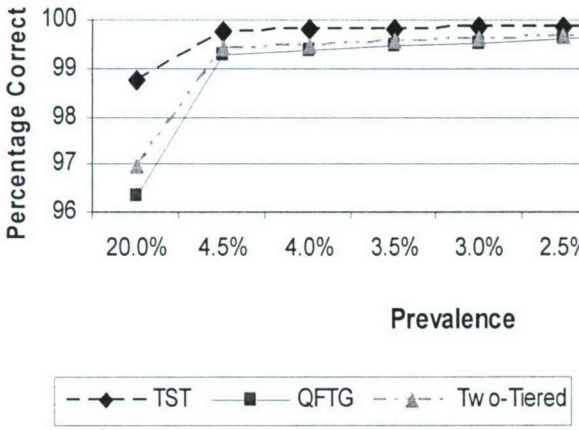

Figure 2. *Negative Predictive values of the TST, QFTG, and a two-tiered test.*

From these predictive values, it is determined that the test presenting the least amount of risk for use in the government is the TST. Although the two-tiered screening and QFTG have a higher positive predictive value, this likely due to the decreased sensitivity of QFTG when compared to the TST. The decreased sensitivity results in fewer true positive patients and false positive patients being identified. The effect of lowering of the positive sample pool increases the odds of those being in the sample as true positives, thus increasing the positive predictive value. However, researchers caution this may not be the case (Mazurek et al, unpublished). The ability of the TST to identify a person with TB as positive will reduce the threat of the disease spreading, thus reduce the threat to military readiness. This increased sensitivity when using a 10 mm cut-point allows for this to occur.

Table 3 is the 2x2 epidemiology table computed for the Mantoux TST scenario, or otherwise known as business as usual. A sample size of 37,344 has been used because it is the difference from the 37,500 RTC projected population, minus the 156 who would be treated based on their history alone taken from historical treatment figures. From the subgroup of 1,830 positive tests, 156 individuals were added based upon historical data to total 1,986 recruits that would be treated for LTBI using this scenario.

Table 3

Mantoux TST 2x2 table with 95% sensitivity and 99% specificity.

		Disease		
		Present	Absent	
Test	Positive	1,472	358	1,830
	Negative	77	35,436	35,514
		1,550	35,794	37,344

Table 4 is the 2x2 table computation built around the scenario of utilizing the QFTG test at RTC for all TB screenings. The same sample size of 37,344 has been applied to the QFTG characteristics. A marked decline of 286 false positive results is due to the increased specificity of the QFTG. Caution needs to be taken here though as the number of false-negative results increases by 155 (232 false negative with the QFTG – 77 false negatives with the Mantoux TST) when compared to the TST.

Table 4

QFTG 2x2 table with 85% sensitivity and 99.8% specificity.

		Disease		
		Present	Absent	
Test	Positive	1,317	72	1,389
	Negative	232	35,723	35,955
		1,549	35,795	37,344

Table 5 is the last 2x2 table developed based on the scenario of a two-tiered testing model. The goal of this model was to reduce the number of recruits who would undergo the more expensive QFTG test, while still applying a screening methodology that would accurately identify those with TB or LTBI. The two-tier testing process is the way HIV testing is done in the military. The first test for HIV is an ELISA test, similar to the QFTG. Positive results then undergo a more expensive and accurate western-blot test in an attempt to minimize false positive and false negative results. The sensitivity of the two-tiered model where all TSTs reading 10 mm or more would undergo a confirmatory QFTG test has been calculated to be 87.5%, with the specificity being 99.99%. The CDC (Mazurek et al. 2005) guidelines for using QuantiFERON®-TB Gold does not recommend this two-tiered or confirmatory testing model. The reasons are not stated, but it is speculated that this was done to eliminate concerns over which test is the

definitive standard for TB testing. If one test was endorsed as the confirmatory test over the other, many serial TB screening programs would feel the need to switch to that test or utilize the two-tiered testing model. This would increase costs to many programs that are constrained by tight budgets. The stance taken by the CDC eliminates the expense increase with their recommendation. By doing so, it allows Public Health Departments and other screening programs to choose a test that best fits their needs without having to worry about using a sub-standard test or face cost-overruns. However, because of the potential impact on readiness, it is added in this analysis.

Table 5

Two tier testing 2x2 table. QFTG testing after TST >5 mm.

		Disease		
(QFTB after Positive TST)		Present	Absent	
Test	Positive	1,303	1	1,304
	Negative	169	357	526
		1,472	358	

When comparing the results of the three 2x2 tables, one must consider not only those who test positive and negative, but the number of false positives and false negative results as well. False results undermine the intent of the screening program by overlooking those who are diseased but not treated, and add cost to the total screening program by treating those who do not have the disease. The false negatives lead to increased risk for the population when a communicable disease is screened for, while false positives add risks to patients who are treated without truly having the disease.

The results indicate the QFTG and the two-tiered process are more discriminate when identifying false positives. This will significantly reduce workload in the BHC's. Again

however, the false negatives go up dramatically when using the two tiered process. The most alarming factor is the total number of false negatives when using the two-tiered scenario includes the 77 subjects missed by the TST, and adds the 169 subjects missed by the confirmatory QFTG. This potential number of LTBI cases missed now totals 246, a significant increase from the stand alone TST testing method. The exact measure of risk incurred by the government from switching tests is something that cannot be addressed however because of numerous variables with wide-ranging estimates, i.e. LTBI to TB conversion rate, LTBI transmission rate, and TB transmission rate. However, the risk is certainly higher when using tests other than the TST.

The cost for each program is the third part of the analysis that will be used in the decision matrix. Medical decisions are not made wholly from financial data, but costs must be measured for planning purposes whenever changes in business practices are being considered. By again utilizing the data extracted from the 2x2 tables, a cost value has been assigned to each part of the tuberculosis control program. These are listed in Table 6.

The costs for isoniazid therapy include the costs from nine months of medication plus a $5.72 fee each time the prescription or refill is completed. The cost for follow-up visits was determined by the cost of operating the clinic, minus ancillary supplies, and then divided by the clinical workload in the RTC clinic. Each visit to BHC 1007 was determined to cost $60. This was applied to all initial visits and follow-up visits. Radiology fees were determined by totaling the departments' overhead costs minus the cost of x-ray film, then dividing that number by the total number of procedures. Individual film prices for the two separate chest x-rays were then added to this expense which totaled $69.36 per patient.

The cost of the Mantoux TST was more difficult to determine. The syringe and PPD solution are inexpensive supplies that must be used for each test. The most expensive part of the test would be the military manpower needed to draw, plant, read and then record each test result.

Other studies were looked at in an effort to compare their Mantoux TST serial testing costs to RTC cost calculations. Their use was rejected when the range of costs was discovered to be from $40 to $300. This was too large of a difference for a realistic estimate. Instead the Champus Maximum Allowable Charge was inserted, which for the northern Illinois area is $11.60.

Table 6

Cost comparison of the 3 separate testing scenarios.

	Mantoux TST	QFTG for all	Two-tier testing
TSTs Given	37,344	0	37,344
Positive Testers	1,778	1,341	1,304
Positive from History	156	156	156
INH Therapy	1,934	1,497	1,460
Cost of PPD ($11.60/test)	$433,190	$0	$433,190
Cost of INH ($61.88/patient)	$119,676	$92,634	$90,345
Cost of x-ray ($69.36/exam)	$134,142	$103,832	$101,266
Cost of F/U ($480 for 8 visits)	$928,320	$718,560	$700,800
Cost of QFTG ($35/test)	$0	$1,312,500	$51,100
Total Costs	$1,615,329	$2,227,526	$1,376,701
Difference from BAU	$0	($612,198)	**$238,628**

The cost of testing with QuantiFERON®-TB Gold is based on a quote from the University of Illinois at Chicago (UIC) laboratory. Included in these costs were all supplies, the pick-up and delivery of samples, and returning the results to RTC Great Lakes. UIC is the only laboratory approved by Celestis to use the QuantiFERON®-TB Gold product in this region. Efforts were made to determine costs within the pathology facilities at Great Lakes and resulted in a test cost $15.00 cost without adding the cost of additional personnel to operate the test.

Estimates of one enlisted tech operating the approximate 37,500 QFTG tests add an additional cost of $1.50 to the test. If only 2,000 tests were conducted as in the two-tiered test scenario, the added cost per test would be $27.00. This large range of manpower costs is not useful. Alternatively, the current laboratory manpower cost of $13.25 per test as calculated from EAS-IV data was added to the $15.00 test cost to total $28.25 per QFTG test. That is cheaper than the UIC pathology quote; however, the $35 cost will be used in the calculations since it is a known variable and thus a more reliable figure.

Discussion

Utilizing the data from the previous tables and figures, a decision matrix was constructed ranking each test in regards to accuracy, government risk, patient risk, cost and ease of process. Although the ease of process has not been addressed in the method section, it is an important consideration. It is assumed that a single test is easier to operate than a two-tiered test. The other assumption is that a test requiring only one patient visit is an easier process for the patients as well as the staff to coordinate. Using those assumptions, the QFTG test will be designated as the first test of choice. The TST has the second easiest process due to the fact that the process is already in place. The other consideration is that with a two-tiered process, you have all the attributes of the two individual process rolled into one. This adds inefficiencies that this analysis would like to avoid. Therefore, the two-tiered process is ranked third.

The test accuracy variable was determined by adding the true positive and true negative results as these are the accurate test numbers. Table 7 has been added to summarize each 2x2 table results. The two-tiered methodology is demonstrably the most accurate way of identifying correctly those who do not have the disease and thus the best test in regards to accuracy. The second most accurate test is QFTG; followed by the TST in third.

Table 7

Summary of 2x2 tables.

	TST	QFTG	Two-tier
True Positive	1,472	1,317	1,303
True Negative	35,436	35,723	35,793
Total True	36,908	37,040	37,096
False Positive	358	72	1
False Negative	77	232	246
Total False	435	304	247

Government risk is not measured the same way it is for a private provider because the Ferres Doctrine eliminates the threat of lawsuit for malpractice. The main concerns of government risk are military readiness. Because the TST is more inclusive and has the highest negative predictive value when identifying those with TB, it is the best test based on reducing government risk. Again this is attributable to the 10 mm cut-point that is used at RTC. The distant second place scenario is then the two-tiered method.

Finally, the cost analysis is considered in the decision matrix. Table 6 demonstrates the cost of each tuberculosis screening exam, as well as the treatment costs associated with a recruit being enrolled in the tuberculosis surveillance program. The largest cost in each of the scenarios is the price of the follow-up visits which are required when a person is identified as having LTBI. Each person entered into the tuberculosis surveillance program costs approximately $620 to the government. In order to save money, the test must correctly identify those with the disease to prevent a TB outbreak and the expense that goes with an outbreak investigation; as well as correctly identifies those without it, to reduce costs of unnecessary treatments. The least costly way to screen for TB was the two-tiered method. The cost reduction here was from limiting the number of people enrolled in the TB surveillance program. The increase risk that comes with the

costs reduction as seen in the two-tiered scenario is difficult to project. Because of the risk, serious consideration needs to be made when cost factors alone are factored into the decision matrix. If 300 sailors contracted LTBI and then needed to be treated because of a faulty test process, the savings of $168,000 from that process would be eliminated by the $186, 000 cost of treating them.

Table 8

Decision matrix for the tables 1 through 9.

Evaluation Criteria	*Metric*	*Importance*	TST	QFTG	Two-tiered
Government Risk	NPV	1	1	3	2
Accuracy	Correct Results	2	3	2	1
Process Ease	Numb. of visits	3	2	1	3
Cost	Overall cost of test	4	2	3	1
Score			8	9	7

Note: Rank of importance was tabulated with 1 denoting the highest importance for test consideration, while 3 was the lowest importance for the metric.

The results of the decision matrix indicate a close score when comparing the TST to the two-tier methodology. Each criterion's rank of importance was next used to determine if this matrix was valid, determining which test would actually be better for the Navy and not just RTC. From the rankings given, the apparent best test method and process at RTC is the two-tiered testing process. However, if the rankings of cost were eliminated which could develop because of the increased likelihood of a TB outbreak, the TST equal the score of the two-tier process.

Recommendations

The answer to the question, "Is there a difference in the likely outcomes of screening test when comparing the Mantoux TST to the QuantiFERON®-TB Gold?" is yes. The increased specificity obtained by utilizing a two-tiered testing process allows for a more accurate screening method in identifying those in RTC who do not have a form of *Mycobacterium tuberculosis* infection. However, the improvement in the identification of true negatives does not meet the needs of the primary military mission. It is the utility of the TST which allows military medicine officials to alter the cut-point, thus making this test more inclusive and reducing the risk to readiness.

The increased specificity that is apparent with the QFTG appears to be useful if there is a concern with a TST result that is between 5 and 9 mms. These are the people who may be "boosted" in the civilian environment, but because of time restraints and the immunization schedule, this does not occur. A quick blood test would give another piece of information to the physician and help them decide what course of therapy should be taken, while still allowing for the seven immunizations to be given to the patient during the same visit.

The process advantages of having a one visit test for TB may offer savings that were not considered in this analysis. Medical facilities that have a low TST reading rate could prove that a test with a known result is better than not knowing the test result. At RTC this is not a problem, as all recruits are marched in together for their results to be read, and the rate of completed TST results exceeds 99%.

Other scenarios were investigated to further reduce costs or estimate risk but were not included within this analysis because of their lack of scientific vigor. There are many situations where a physician may want another test performed to help them with the decision process. Within some restrictive scenario's they should be allowed the freedom to use whatever test they

deem necessary. The CDC (Mazurek et al., 2005) clearly states that if there is a positive TST, there is no reason to complete a QFTG. However, because of the inclusive cut-point used at RTC a physician may feel that another test is a responsible expenditure of $35 instead of the $620 cost of treating them for nine months. This is especially true when a person has no known risk factors for TB.

This analysis only looked into the mass screening process at Recruit Training Command, Great Lakes, IL. The determination of how the test would impact results at other installations, including cost effects, is dependent upon their individual circumstances including their historical TB prevalence and percentage of TSTs that are read.

The debate will likely continue concerning which test is best for TB screening. The research does not demonstrate a single superior test at this time. Some researches believe that the TST false positive rate is too high, and the QFTG is more accurate. Others believe that the QFTG does not identify LTBI carriers as reliably as the TST the likely will not be resolved soon. However, the information gained from conducting a long term two-tiered test process may provide answers to these and other questions.

With the criteria established utilizing current research for the Mantoux TST and QuantiFERON®-TB Gold tests, this analysis indicates that two-tiered testing may be cost effective, as well as a way to eliminate false negative results. However, military priorities would likely indicate that the increased risk from changing testing scenarios would not be prudent. Other tests are in the works based upon the same technology as QuantiFERON®-TB Gold. Their usefulness should be analyzed if they are approved in the future by the FDA.

References

Bass, J. B., Jr. (2003). How good is the tuberculin skin test? *Infection Control & Hospital Epidemiology, 24*(11), 797-798.

Bass, J. B., Jr. (2001).The tuberculin skin test. In M.J. Field (Ed.), *Tuberculosis in the Workplace* (pp. B1-B10). Washington, DC: National Academy Press.

Center for Disease Control (1997). *Case definitions for infectious conditions under public health surveillance. Morbidity and Mortality Weekly Report, 46*, 40-41.

Dunlap, N. E., Bass, J. B., Fujiwara, P., Hopewell, P., Horsburgh, C. R, Salfinger, M. et al. (2000). Diagnostic standards and classification of tuberculosis in adults and children. *American Journal of Respiratory and Critical Care Medicine*, *161,* 1376-1395.

Ferrara, G., Losi, M., Meacci, M., Meccugni, B., Piro, R., Roversi, P., et al. (2005). Routine hospital use of a new commercial whole blood interferon-γ assay for the diagnosis of tuberculin infection. *American Journal of Respiratory and Critical Care Medicine*, *172,* 631-635.

Huebner, R .E., Schein, M. F., & Bass, J.B. (1993). The tuberculin skin test. *Clinical Infectious Diseases, 17,* 968 – 975.

Jensen, P. A., Lambert, L. A., Iademarco, M. F., & Ridzon, R. (2005). Guidelines for preventing the transmission of *Mycobacterium tuberculosis* in health-care settings. *Morbidity and Mortality Weekly Report, 54(RR17), 1-141.*

Institute of Medicine (2000). *Ending neglect: The elimination of tuberculosis in the United States.* Washington, DC: National Academy Press.

Jekel, J. F., Katz, D.L., & Elmore, J.G. (2001). *Epidemiology, biostatistics, and preventive medicine* (2nd ed.) Philadelphia: W.B. Saunders.

Kang, Y., Lee, H., Yoon, H., Cho, B., Han, S., Shim, Y., & Yim, J. (2005). Discrepancy between the tuberculin skin test and the whole-blood interferon γ assay for the diagnosis of latent tuberculosis infection in an intermediate tuberculosis-burden country. *JAMA, 293(22),* 2756-2761.

Kendig, E., Kirkpatrick, B.V., Carter, W., Hill, F.A., Caldwell, K. & Entwistle, M. (1998). Underreading of the tuberculin skin test reaction. *Chest, 113*(5), 1175-1177.

Lamar, J. E. & Malakooti, M. A. (2003). Tuberculosis-outbreak investigation of a U.S. Navy amphibious ship crew and the Marine expeditionary unit aboard, 1998. *Military Medicine*, *168*(7), 523-527.

Mazurek, G. H., Zajdowicz, M. J., Hankinson, A. L., Costigan, D., Rothel, J. S., Toney, S. R., Robles, B., Van Herpen, J., Wild, S., Daniels, L. J., LoBue, P., Pascual, B., Shang, N., & Keep, L. W. (2004) Detection of *Mycobacterium tuberculosis* infection in United States Navy Recruits using the tuberculin skin test or whole blood interferon-gamma release assays. Unpublished study data.

Mazurek, G. H., Jereb, J., LoBue, P., Iademarco, M. F., Metchock, B. & Vernon, A. (2005). Guidelines for using the QuantiFERON®-TB Gold test for detecting mycobacterium tuberculosis infection, United States. *Morbidity & Mortality Weekly Report, 54,* 49-55.

Mori, T., Sakatani, M., Yamagishi, F., Takashima, T., Kawabe, Y., & Nagao, K. et al. (2004). Specific detection of tuberculosis infection: and interferon-γ based assay using new antigens. *American Journal of Respiratory and Critical Care Medicine*, *170*, 59-64.

Nguyen, M., Perry, S., & Parsonnet, J. (2005). QuantiFERON®-TB Gold predicts tuberculin skin test boosting in U.S. foreign-born. *International Journal of Tuberculosis and Lung Disease, 9(9),* 985-991.

Petitti, D. (2000). M*eta-analysis, decision-analysis, and cost-effective analysis: Methods for quantitative synthesis in medicine (2nd ed.)* New York, Oxford University Press.

Smith, B., Ryan, M. A., Gray, G.C., Polonsky, J. M., & Trump, D. H. (2002) Tuberculosis infection among young adults enlisting in the United States Navy. *International Journal of Epidemiology*; *31*, 934-939.

Smith, L. & Schillaci, R. (1987) Tuberculosis Control in the Military Population: Apathy Revisited. *Military Medicine Journal, 152*(8) p. 421-423

Taylor, Z., Nolan, C., & Blumberg, H. (2005). Controlling tuberculosis in the United States: Recommendations from the American Thoracic Society, CDC, and Infectious Diseases Society of America. *Morbidity and Mortality Weekly Report, 54(RR12),* 1-81.

Whalen, C. C. (2005). Diagnosis of latent tuberculosis infection: Measure for measure. *JAMA*, *293*(22), 2785-287.

World Health Organization (2005, April). Tuberculosis fact sheet. Retrieved Sep. 17, 2005 at http://www.who.int/mediacentre/factsheets/fs104/en/index.html

Appendix A: Abbreviations

BHC – branch health clinic

CDC – Center for Disease Control

FDA – Food and Drug Administration

LTBI - latent tuberculosis infection

MMWR – Morbidity & Mortality Weekly Report

NPV – negative predictive value

NTM – non-tuberculosis mycobacterium

PPD – purified protein derivative

PPV – positive predictive value

QFTG – QuantiFERON®-TB Gold

RT – Recruit Training

RTC – recruit training command

TB – tuberculosis

TST – tuberculin skin test

UIC – University of Illinois at Chicago

Appendix B. 2x2 table summary

4.0% Prevalence TST 95% Sensitive, 99% Specific

Table 5.

2x2 table with 95% sensitivity and 99% specificity.

Mantoux TST		Disease		
		Present	Absent	
Test	Positive	1472	358	1,830
	Negative	77	35,436	35,514
		1,550	35,794	

4.0% Prevalence QFTG 85% Sensitive, 99.8% Specific

Table 6.

2x2 table with 85% sensitivity and 99.8% specificity.

QuantiFERON TB Gold		Disease		
		Present	Absent	
Test	Positive	1,317	72	1,389
	Negative	232	35,723	35,955
		1,549	35,850	

83% Prevalence after Positive TST.

Table 7.

2x2 table 2nd tier of QFTG test post TST.

Two-tiered Test QFTB after Mantoux Positive		Disease		
		Present	Absent	
Test	Positive	1303	1	1304
	Negative	169	358	526
		1,472	359	1778

*83% Prevalence based on PPV of TST. 1778

Sensitivity and specificity of QFTG remain the same.